

SEQ ID NO:120

ycfE protein - Escherichia coli (strain K-12) cryptic prophage e14
 C;Species: Escherichia coli
 C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 C;Accession: B64861; S07185
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.; Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of Escherichia coli K-12.
 A;Reference number: A64720; MUID:97426617; PMID:9278503
 A;Accession: B64861
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-179 <BLAT>
 A;Cross-references: GB:AE000214; GB:U00096; NID:g1787382; PIDN:AAC74241.1; PID:g1787403; UWGP:b1157
 A;Experimental source: strain K-12, substrain MG1655
 R;Plasterk, R.H.A.; van de Putte, P.
 EMBO J. 4, 237-242, 1985
 A;Title: The invertible P-DNA segment in the chromosome of Escherichia coli.
 A;Reference number: S07185; MUID:85257443; PMID:3894006
 A;Accession: S07185
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 'IKSDGTVQTALE',13,'L',15-164,'AVRVILHLTIS' <PLA>
 A;Cross-references: EMBL:X01805
 C;Genetics:
 A;Gene: ycfE
 A;Map position: 25 min
 A;Genome: cryptic prophage e14
 C;Superfamily: phage P1 major tail fiber protein S

Query Match 1.6%; Score 8; DB 2; Length 179;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;

Qy 20 GLGEGSAL 27
 |||||
 Db 15 GLGEGSAL 22

ID AAB43919 standard; Protein; 602 AA.
 XX
 AC AAB43919;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated protein sequence SEQ ID NO:1364.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
 KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX
 OS Homo sapiens.
 XX
 PN WO200055350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05882.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-587533/55.
 DR N-PSDB; AAC78128.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer -
 XX
 PS Claim 11; Page 2024-2026; 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnerary; immunomodulator;
 CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate

CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.

XX

SQ Sequence 602 AA;

Query Match 99.8%; Score 2816; DB 21; Length 602;
Best Local Similarity 99.8%; Pred. No. 2e-253;
Matches 513; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

| | | | | |
|----|-----|---|---------------------------------------|-----|
| Qy | 1 | MSWPRRLLLRYLFPALLLHGLGEGSALLHPDSRSHPRSLEKSAWRAFKE | SQCHHMLKHLH | 60 |
| | | | | |
| Db | 89 | MSWPRRLLLRYLFPALLLHGLGEGSALLHPDSRSHPRSLEKSAWRAFKE | SQCHHMLKHLH | 148 |
| Qy | 61 | NGARITVQMPPTIEGHVWSTGCEVRSGPEFITRSYRFYHNNTFKAYQFY | YGSNRCTNPTY | 120 |
| | | | | |
| Db | 149 | NGARITVQMPPTIEGHVWSTGCEVRSGPEFITRSYRFYHNNTFKAYQFY | YGSNRCTNPTY | 208 |
| Qy | 121 | TLIIRGKIRLRQASWIIRGGTEADYQLHNVQVICHTEAVA | EKLGGQVNRCTCPGFLADGGP | 180 |
| | | | | |
| Db | 209 | TLIIRGKIRLRQASWIIRGGTEADYQLHNVQVICHTEAVA | EKLGGQVNRCTCPGFLADGGP | 268 |
| Qy | 181 | WVQDVAYDLWREENGCECTKAVNFAMHELQLIRVEKQYLHHNLDHL | VEELFLGDIHTDAT | 240 |
| | | | | |
| Db | 269 | WVQDVAYDLWREENGCECTKAVNFAMHELQLIRVEKQYLHHNLDHL | VEELFLGDIHTDAT | 328 |
| Qy | 241 | QRMFYRPSSYQPPLQNAKNDHACIACRIIYRSDEHHPPILPPKADLT | IGLHGEWVSQRC | 300 |
| | | | | |
| Db | 329 | QRMFYRPSSYQPPLQNAKNDHACIACXIIYRSDEHHPPILPPKADLT | IGLHGEWVSQRC | 388 |
| Qy | 301 | EVRPEVLFLTRHFIFHDNNNTWEGHYHYSDPVCKHPTFSIYARGRY | SRGVLSSRVMGGT | 360 |
| | | | | |
| Db | 389 | EVRPEVLFLTRHFIFHDNNNTWEGHYHYSDPVCKHPTFSIYARGRY | SRGVLSSRVMGGT | 448 |
| Qy | 361 | EFVFKVNHMKVTPMDAATASLLNVFNGNECGAEGSWQVGIQQDV | THTNGCVALGIKLPH | 420 |
| | | | | |
| Db | 449 | EFVFKVNHMKVTPMDAATASLLNVFNGNECGAEGSWQVGIQQDV | THTNGCVALGIKLPH | 508 |
| Qy | 421 | EYEIFKMEQDARGRYLLFNGQRP | SDGSSPDRPEKRATSYQMPLVQCASSSPRAEDLAEDS | 480 |
| | | | | |
| Db | 509 | EYEIFKMEQDARGRYLLFNGQRP | SDGSSPDRPEKRATSYQMPLVQCASSSPRAEDLAEDS | 568 |
| Qy | 481 | GSSLYGRAPGRHTWSLLLAALACLVPLLHWNIRR | | 514 |
| | | | | |
| Db | 569 | GSSLYGRAPGRHTWSLLLAALACLVPLLHWNIRR | | 602 |

ID AAC78128 standard; cDNA; 2162 BP.
 XX
 AC AAC78128;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated gene sequence SEQ ID NO:522.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
 KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200055350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05882.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-587533/55.
 DR P-PSDB; AAB43919.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer -
 XX
 PS Claim 1; Page 1051-1052; 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnerary; immunomodulator;
 CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate

| | | | |
|----|------|--|------|
| Qy | 201 | AlaValAsnPheAlaMetHisGluLeuGlnLeuIleArgValGluLysGlnTyrLeuHis | 220 |
| Db | 866 | | |
| | | 866 GCCGTGAACCTTTGCCATGCATGAACTTCAGCTCATCCGGGTGGAGAAGCAGTACCTTCAC | 925 |
| Qy | 221 | HisAsnLeuAspHisLeuValGluGluLeuPheLeuGlyAspIleHisThrAspAlaThr | 240 |
| Db | 926 | | |
| | | 926 CACAACCTCGACCACCTGGTCGAGGAGCTCTTCCTTGGTGACATTCACACTGATGCCACC | 985 |
| Qy | 241 | GlnArgMetPheTyrArgProSerSerTyrGlnProProLeuGlnAsnAlaLysAsnHis | 260 |
| Db | 986 | | |
| | | 986 CAGAGGATGTTCTACCGGCCCTCCAGTTACCAGCCCCCTCTGCAGAATGCCAAGAACCAC | 1045 |
| Qy | 261 | AspHisAlaCysIleAlaCysArg-IleIleTyrArgSerAspGluHisHisProProIl | 280 |
| Db | 1046 | | |
| | | 1046 GACCATGCCTGCATCGCCTGTSG-GATCATCTATCGGTCAGACGAGCACCACCCTCCCAT | 1104 |
| Qy | 280 | eLeuProProLysAlaAspLeuThrIleGlyLeuHisGlyGluTrpValSerGlnArgCy | 300 |
| Db | 1105 | | |
| | | 1105 CCTGCCCCCAAAGGCAGACCTGACCATCGGCCTGCACGGGGAGTGGGTGAGCCAGCGCTG | 1164 |
| Qy | 300 | sGluValArgProGluValLeuPheLeuThrArgHisPheIlePheHisAspAsnAsnAs | 320 |
| Db | 1165 | | |
| | | 1165 TGAGGTGCGCCCCGAAGTCCTCTTCTCACC CGCCACTTCATCTTCCATGACAACAACAA | 1224 |
| Qy | 320 | nThrTrpGluGlyHisTyrTyrHisTyrSerAspProValCysLysHisProThrPheSe | 340 |
| Db | 1225 | | |
| | | 1225 CACCTGGGAGGGCCACTACTACCACTACTCAGACCCGGTGTGCAAGCACCCACCTTCTC | 1284 |
| Qy | 340 | rIleTyrAlaArgGlyArgTyrSerArgGlyValLeuSerSerArgValMetGlyGlyTh | 360 |
| Db | 1285 | | |
| | | 1285 CATCTACGCCCGGGGCCGCTACAGCCGCGGCGTCTCTCGTCCAGGGTCATGGGAGGCAC | 1344 |
| Qy | 360 | rGluPheValPheLysValAsnHisMetLysValThrProMetAspAlaAlaThrAlaSe | 380 |
| Db | 1345 | | |
| | | 1345 CGAGTTCGTGTTCAAAGTGAATCACATGAAGGTCACCCCATGGATGCGGCCACAGCCTC | 1404 |
| Qy | 380 | rLeuLeuAsnValPheAsnGlyAsnGluCysGlyAlaGluGlySerTrpGlnValGlyIl | 400 |
| Db | 1405 | | |
| | | 1405 ACTGCTCAACGTCTTCAACGGGAATGAGTGCGGGGCCGAGGGCTCCTGGCAGGTGGGCAT | 1464 |
| Qy | 400 | eGlnGlnAspValThrHisThrAsnGlyCysValAlaLeuGlyIleLysLeuProHisTh | 420 |
| Db | 1465 | | |
| | | 1465 CCAGCAGGATGTGACCCACACCAATGGCTGCGTGGCCCTGGGCATCAAACCTACCTCACAC | 1524 |
| Qy | 420 | rGluTyrGluIlePheLysMetGluGlnAspAlaArgGlyArgTyrLeuLeuPheAsnGl | 440 |
| Db | 1525 | | |
| | | 1525 GGAGTACGAGATCTTCAAATGGAACAGGATGCCCGGGGGCGCTATCTGCTGTTCAACGG | 1584 |
| Qy | 440 | yGlnArgProSerAspGlySerSerProAspArgProGluLysArgAlaThrSerTyrGl | 460 |
| Db | 1585 | | |
| | | 1585 TCAGAGGCCAGCGACGGGTCCAGCCAGACAGGCCAGAGAAGAGAGCCACGTCTACCA | 1644 |
| Qy | 460 | nMetProLeuValGlnCysAlaSerSerSerProArgAlaGluAspLeuAlaGluAspSe | 480 |
| Db | 1645 | | |
| | | 1645 GATGCCCTTGGTCCAGTGTGCCTCTCTTCGCCGAGGGCAGAGGACCTYGCAGAAGACAG | 1704 |

Qy 480 rGlySerSerLeuTyrGlyArgAlaProGlyArgHisThrTrpSerLeuLeuLeuAlaAl 500
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1705 TGGAAGCAGCCTGTATGGCCGGGCCCCCTGGGAGGCACACCTGGTCCCTGCTGCTGGCTGC 1764

Qy 500 aLeuAlaCysLeuValProLeuLeuHisTrpAsnIleArgArg 514
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1765 ACTTGCTGYCTTGTCCTCTGCTGCATTGGAACATCCGCAGA 1807

ID AAC78128 standard; cDNA; 2162 BP.
 XX
 AC AAC78128;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated gene sequence SEQ ID NO:522.
 XX
 PN WO200055350-A1.
 XX
 PS Claim 1; Page 1051-1052; 2352pp; English.
 XX
 SQ Sequence 2162 BP; 462 A; 683 C; 594 G; 412 T; 11 other;

Query Match 85.3%; Score 1583; DB 21; Length 2162;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1833; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 19 GCGGAGGGCAGAGCGCGCGCCAGTTGCCCGGGCACCAAATCGGAGCGCGGCGTGCGGGA 78
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 185 GCGGAGGGCAGAGCGCGCGCCAGTTGCCCGGGCACCAAATCGGAGCGCGGCGTGCGGGA 244

Qy 79 GGGCCCAGAGCAGGACTGGAAATGTCCTGGCCGCGCGCCTCCTGCTCAGATACCTGTTC 138
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 245 GGGCCCAGAGCAGGACTGGAAATGTCCTGGCCGCGCGCCTCCTGCTCAGATACCTGTTC 304

Qy 139 CCGGCCCTCCTGCTTCACGGGCTGGGAGAGGGTTCTGCCCTCCTTCATCCAGACAGCAGG 198
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 305 CCGGCCCTCCTGCTTCACGGGCTGGGAGAGGGTTCTGCCCTCCTTCATCCAGACAGCAGG 364

Qy 199 TCTCATCCTAGGTCCTTAGAGAAAAGTGCCTGGAGGGCTTTTAAGGAGTCACAGTGCCAT 258
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 365 TCTCATCCTAGGTCCTTAGAGAAAAGTGCCTGGAGGGCTTTTAAGGAGTCACAGTGCCAT 424

Qy 259 CACATGCTCAAACATCTCCACAATGGTGCAAGGATCACAGTGCAGATGCCACCTACAATC 318
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 425 CACATGCTCAAACATCTCCACAATGGTGCAAGGATCACAGTGCAGATGCCACCTACAATC 484

Qy 319 GAGGGCCACTGGGTCTCCACAGGCTGTGAAGTAAGGTCAGGCCCAGAGTTCATCACAAGG 378
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 485 GAGGGCCACTGGGTCTCCACAGGCTGTGAAGTAAGGTCAGGCCCAGAGTTCATCACAAGG 544

Qy 379 TCCTACAGATTCTACCACAATAACACCTTCAAGGCCTACCAATTTTATTATGGCAGCAAC 438
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 545 TCCTACAGATTCTACCACAATAACACCTTCAAGGCCTACCAATTTTATTATGGCAGCAAC 604

| | | | |
|----|------|---|------|
| Qy | 439 | CGGTGCACAAATCCCACTTATACTCTCATCATCCGGGGCAAGATCCGCCTCCGCCAGGCC | 498 |
| Db | 605 | CGGTGCACAAATCCCACTTATACTCTCATCATCCGGGGCAAGATCCGCCTCCGCCAGGCC | 664 |
| Qy | 499 | TCCTGGATCATCCGAGGGGGCACGGAAGCCGACTACCAGCTGCACAACGTCCAGGTGATC | 558 |
| Db | 665 | TCCTGGATCATCCGAGGGGGCACGGAAGCCGACTACCAGCTGCACAACGTCCAGGTGATC | 724 |
| Qy | 559 | TGCCACACAGAGGCGGTGGCCGAGAAGCTGGGCCAGCAGGTGAACCGCACATGCCCCGGC | 618 |
| Db | 725 | TGCCACACAGAGGCGGTGGCCGAGAAGCTCGGCCAGCAGGTGAACCGCACATGCCCCGGC | 784 |
| Qy | 619 | TTCTTCGCAGACGGGGGTCCCTGGGTGCAGGACGTGGCCTATGACCTCTGGCGAGAGGAG | 678 |
| Db | 785 | TTCTTCGCAGACGGGGGTCCCTGGGTGCAGGACGTGGCCTATGACCTCTGGCGAGAGGAG | 844 |
| Qy | 679 | AACGGCTGTGAGTGCACCAAGGCCGTGAACTTTGCCATGCATGAACTTCAGCTCATCCGG | 738 |
| Db | 845 | AACGGCTGTGAGTGCACCAAGGCCGTGAACTTTGCCATGCATGAACTTCAGCTCATCCGG | 904 |
| Qy | 739 | GTGGAGAAGCAGTACCTTCACCACAACCTCGACCACCTGGTCGAGGAGCTCTTCCTTGGT | 798 |
| Db | 905 | GTGGAGAAGCAGTACCTTCACCACAACCTCGACCACCTGGTCGAGGAGCTCTTCCTTGGT | 964 |
| Qy | 799 | GACATTACACTGATGCCACCCAGAGGATGTTCTACCGGCCCTCCAGTTACCAGCCCCCT | 858 |
| Db | 965 | GACATTACACTGATGCCACCCAGAGGATGTTCTACCGGCCCTCCAGTTACCAGCCCCCT | 1024 |
| Qy | 859 | CTGCAGAATGCCAAGAACCACGACCATGCCTGCATCGCCTGTSGGATCATCTATCGGTCA | 918 |
| Db | 1025 | CTGCAGAATGCCAAGAACCACGACCATGCCTGCATCGCCTGTSGGATCATCTATCGGTCA | 1084 |
| Qy | 919 | GACGAGCACCACCCTCCCATCCTGCCCCCAAAGGCAGACCTGACCATCGGCCTGCACGGG | 978 |
| Db | 1085 | GACGAGCACCACCCTCCCATCCTGCCCCCAAAGGCAGACCTGACCATCGGCCTGCACGGG | 1144 |
| Qy | 979 | GAGTGGGTGAGCCAGCGCTGTGAGGTGCGCCCCGAAGTCCTCTTCCTCACCCGCCACTTC | 1038 |
| Db | 1145 | GAGTGGGTGAGCCAGCGCTGTGAGGTGCGCCCCGAAGTCCTCTTCCTCACCCGCCACTTC | 1204 |
| Qy | 1039 | ATCTTCCATGACAACAACAACACCTGGGAGGGCCACTACTACCACTACTCAGACCCGGTG | 1098 |
| Db | 1205 | ATCTTCCATGACAACAACAACACCTGGGAGGGCCACTACTACCACTACTCAGACCCGGTG | 1264 |
| Qy | 1099 | TGCAAGCACCCACCTTCTCCATCTACGCCCCGGGGCCGCTACAGCCGCGGCGTCTCTCG | 1158 |
| Db | 1265 | TGCAAGCACCCACCTTCTCCATCTACGCCCCGGGGCCGCTACAGCCGCGGCGTCTCTCG | 1324 |
| Qy | 1159 | TCCAGGGTCATGGGAGGCACCGAGTTCGTGTTCAAAGTGAATCACATGAAGGTCACCCCC | 1218 |
| Db | 1325 | TCCAGGGTCATGGGAGGCACCGAGTTCGTGTTCAAAGTGAATCACATGAAGGTCACCCCC | 1384 |
| Qy | 1219 | ATGGATGCGGCCACAGCCTCACTGCTCAACGCTCTTCAACGGGAATGAGTGCGGGGCCGAG | 1278 |
| Db | 1385 | ATGGATGCGGCCACAGCCTCACTGCTCAACGCTCTTCAACGGGAATGAGTGCGGGGCCGAG | 1444 |

Qy 1279 GGCTCCTGGCAGGTGGGCATCCAGCAGGATGTGACCCACACCAATGGCTGCGTGGCCCTG 1338
 |||
 Db 1445 GGCTCCTGGCAGGTGGGCATCCAGCAGGATGTGACCCACACCAATGGCTGCGTGGCCCTG 1504
 |||
 Qy 1339 GGCATCAAACCTACCTCACACGGAGTACGAGATCTTCAAAATGGAACAGGATGCCCCGGGGG 1398
 |||
 Db 1505 GGCATCAAACCTACCTCACACGGAGTACGAGATCTTCAAAATGGAACAGGATGCCCCGGGGG 1564
 |||
 Qy 1399 CGCTATCTGCTGTTCAACGGTCAGAGGCCCAGCGACGGGTCCAGCCCAGACAGGCCAGAG 1458
 |||
 Db 1565 CGCTATCTGCTGTTCAACGGTCAGAGGCCCAGCGACGGGTCCAGCCCAGACAGGCCAGAG 1624
 |||
 Qy 1459 AAGAGAGCCACGTCCTACCAGATGCCCTTGGTCCAGTGTGCCTCCTCTTCGCCGAGGGCA 1518
 |||
 Db 1625 AAGAGAGCCACGTCCTACCAGATGCCCTTGGTCCAGTGTGCCTCCTCTTCGCCGAGGGCA 1684
 |||
 Qy 1519 GAGGACCTCGCAGAAGACAGTGAAGCAGCCTGTATGGCCGGGCCCTGGGAGGCACACC 1578
 |||
 Db 1685 GAGGACCTYGCAGAAGACAGTGAAGCAGCCTGTATGGCCGGGCCCTGGGAGGCACACC 1744
 |||
 Qy 1579 TGGTCCCTGCTGCTGGCTGCACTTGCCTGCCTTGTCCCTCTGCTGCATTGGAACATCCGC 1638
 |||
 Db 1745 TGGTCCCTGCTGCTGGCTGCACTTGCCTGYCTTGTCCCTCTGCTGCATTGGAACATCCGC 1804
 |||
 Qy 1639 AGATAGAAGTTTTAGAAAGTTCTATTTTTCCAAACCAGGATTCCTTACTATTGACAGATT 1698
 |||
 Db 1805 AGATAGAAGTTTTAGAAAGTTCTATTTTTCCAAACCAGGATTCCTTACTATTGACAGATT 1864
 |||
 Qy 1699 TTCTTTACCAAAGAAAAGACATTTATTCTTTTGATGCACTTGAATGCCAGAGAACTGTC 1758
 |||
 Db 1865 TKCTTTACCAAAGAAAAGACATTTATTCTTTTGATGCACTTGAATGCCAGAGAACTGTC 1924
 |||
 Qy 1759 CTTCTTTTTCTCCTCTCCCTCCCTCCCAGCCCCCTGAGTCATGAACAGCAAGGAGTGTGTTG 1818
 |||
 Db 1925 CTTCTTTTTCTCCTCTCCCTCCCTCCCAGCCCCCTGAGTCATGAACAGCAAGGAGTGTGTTG 1984
 |||
 Qy 1819 AAGTTTCTGCTTTGAACTCCGTCCAGCCTGATCCCTGG 1856
 |||
 Db 1985 AAGTTTCTGCTTTGAACTCCGTCCAGCCTGATCCCTGG 2022

BM715696

LOCUS BM715696 798 bp mRNA linear EST 28-FEB-2002

DEFINITION UI-E-EJ0-ahj-g-05-0-UI.r2 UI-E-EJ0 Homo sapiens cDNA clone
 UI-E-EJ0-ahj-g-05-0-UI 5', mRNA sequence.

ACCESSION BM715696

VERSION BM715696.1 GI:19028954

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 798)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL **Genome Res. 6 (9), 791-806 (1996)**
MEDLINE 97044477
COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES Location/Qualifiers
 source 1. .798
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-EJ0-ahj-g-05-0-UI"
 /clone_lib="UI-E-EJ0"
 /tissue_type="fetal eyes, lens, eye anterior segment,
 optic nerve, retina, Retina Foveal and Macular, RPE and
 Choroid"
 /dev_stage="fetal and adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-EJ0 is a subtracted cDNA library constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT7T3-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tags for this library are: fetal eyes, AGAATCAAGA
 ; lens, CGATTAGCGA; eye anterior segment, AATGCCGCAT;
 optic nerve, CCATTAAGTG; retina, CCGCG; Retina Foveal and
 Macular, GTCC; RPE and Choroid, ACCTA. This library was
 created for the program, Gene Discovery in the Visual
 System, supported by National Eye Institute (NEI)."
BASE COUNT 184 a 244 c 206 g 164 t
ORIGIN

Alignment Scores:
 Pred. No.: 1.87e-227 Length: 798
 Score: 249.00 Matches: 249
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 48.44% Indels: 0
 DB: 14 Gaps: 0

US-09-720-533-120 (1-514) x BM715696 (1-798)

Qy 10 ArgTyrLeuPheProAlaLeuLeuLeuHisGlyLeuGlyGluGlySerAlaLeuLeuHis 29
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2 AGATACCTGTTCCCGGCCCTCCTGCTTCACGGGCTGGGAGAGGGTTCTGCCCTCCTTCAT 61

 Qy 30 ProAspSerArgSerHisProArgSerLeuGluLysSerAlaTrpArgAlaPheLysGlu 49
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 62 CCAGACAGCAGGTCTCATCCTAGGTCCTTAGAGAAAAGTGCCTGGAGGGCTTTTAAGGAG 121

 Qy 50 SerGlnCysHisHisMetLeuLysHisLeuHisAsnGlyAlaArgIleThrValGlnMet 69
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 122 TCACAGTGCCATCACATGCTCAAACATCTCCACAATGGTGCAAGGATCACAGTGCAGATG 181

 Qy 70 ProProThrIleGluGlyHisTrpValSerThrGlyCysGluValArgSerGlyProGlu 89
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 182 CCACCTACAATCGAGGGCCACTGGGTCTCCACAGGCTGTGAAGTAAGGTCAGGCCCAGAG 241

 Qy 90 PheIleThrArgSerTyrArgPheTyrHisAsnAsnThrPheLysAlaTyrGlnPheTyr 109
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 242 TTCATCACAAGGTCCTACAGATTCTACCACAATAACACCTTCAAGGCCTACCAATTTTAT 301

 Qy 110 TyrGlySerAsnArgCysThrAsnProThrTyrThrLeuIleIleArgGlyLysIleArg 129
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 302 TATGGCAGCAACCGGTGCACAAATCCCACTTATACTCTCATCATCCGGGCAAGATCCGC 361

 Qy 130 LeuArgGlnAlaSerTrpIleIleArgGlyGlyThrGluAlaAspTyrGlnLeuHisAsn 149
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 362 CTCCGCCAGGCCTCCTGGATCATCCGAGGGGGCACGGAAGCCGACTACCAGCTGCACAAC 421

 Qy 150 ValGlnValIleCysHisThrGluAlaValAlaGluLysLeuGlyGlnGlnValAsnArg 169
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 422 GTCCAGGTGATCTGCCACACAGAGGCGGTGGCCGAGAAGCTCGGCCAGCAGGTGAACCGC 481

 Qy 170 ThrCysProGlyPheLeuAlaAspGlyGlyProTrpValGlnAspValAlaTyrAspLeu 189
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 482 ACATGCCCCGGGCTTCCTCGCAGACGGGGTCCCTGGGTGCAGGACGTGGCCTATGACCTC 541

 Qy 190 TrpArgGluGluAsnGlyCysGluCysThrLysAlaValAsnPheAlaMetHisGluLeu 209
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 542 TGGCGAGAGGAGAACGGCTGTGAGTGCACCAAGGCCGTGAACCTTTGCCATGCATGAACCT 601

 Qy 210 GlnLeuIleArgValGluLysGlnTyrLeuHisHisAsnLeuAspHisLeuValGluGlu 229
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 602 CAGCTCATCCGGGTGGAGAAGCAGTACCTTACCACAACCTCGACCACCTGGTCGAGGAG 661

 Qy 230 LeuPheLeuGlyAspIleHisThrAspAlaThrGlnArgMetPheTyrArgProSerSer 249
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 662 CTCTTCCTTGGTGACATTACACTGATGCCACCCAGAGGATGTTCTATCGGCCCTCCAGT 721

 Qy 250 TyrGlnProProLeuGlnAsnAlaLys 258
 ||||||||||||||||||||
 Db 722 TACCAGCCCCCTCTGCAGAATGCCAAG 748

LOCUS BM715696 798 bp mRNA linear EST 28-FEB-2002
 DEFINITION UI-E-EJ0-ahj-g-05-0-UI.r2 UI-E-EJ0 Homo sapiens cDNA clone
 UI-E-EJ0-ahj-g-05-0-UI 5', mRNA sequence.
 ACCESSION BM715696
 VERSION BM715696.1 GI:19028954
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 798)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL **Genome Res.** 6 (9), 791-806 (1996)

BASE COUNT 184 a 244 c 206 g 164 t
 ORIGIN

Query Match 35.4%; Score 657; DB 14; Length 798;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 707; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 126 | CAGATACCTGTTCCCGGCCCTCCTGCTTCACGGGCTGGGAGAGGGTTCTGCCCTCCTTCA | 185 |
| Db | 1 | CAGATACCTGTTCCCGGCCCTCCTGCTTCACGGGCTGGGAGAGGGTTCTGCCCTCCTTCA | 60 |
| Qy | 186 | TCCAGACAGCAGGTCTCATCCTAGGTCCTTAGAGAAAAGTGCCTGGAGGGCTTTTAAGGA | 245 |
| Db | 61 | TCCAGACAGCAGGTCTCATCCTAGGTCCTTAGAGAAAAGTGCCTGGAGGGCTTTTAAGGA | 120 |
| Qy | 246 | GTCACAGTGCCATCACATGCTCAAACATCTCCACAATGGTGCAAGGATCACAGTGCAGAT | 305 |
| Db | 121 | GTCACAGTGCCATCACATGCTCAAACATCTCCACAATGGTGCAAGGATCACAGTGCAGAT | 180 |
| Qy | 306 | GCCACCTACAATCGAGGGCCACTGGGTCTCCACAGGCTGTGAAGTAAGGTCAGGCCCAGA | 365 |
| Db | 181 | GCCACCTACAATCGAGGGCCACTGGGTCTCCACAGGCTGTGAAGTAAGGTCAGGCCCAGA | 240 |
| Qy | 366 | GTTTCATCACAAGGTCTACAGATTCTACCACAATAACACCTTCAAGGCCTACCAATTTTA | 425 |
| Db | 241 | GTTTCATCACAAGGTCTACAGATTCTACCACAATAACACCTTCAAGGCCTACCAATTTTA | 300 |
| Qy | 426 | TTATGGCAGCAACCGGTGCACAAATCCCCTTATACTCTCATCATCCGGGGCAAGATCCG | 485 |
| Db | 301 | TTATGGCAGCAACCGGTGCACAAATCCCCTTATACTCTCATCATCCGGGGCAAGATCCG | 360 |
| Qy | 486 | CCTCCGCCAGGCCTCCTGGATCATCCGAGGGGGCACGGAAGCCGACTACCAGCTGCACAA | 545 |
| Db | 361 | CCTCCGCCAGGCCTCCTGGATCATCCGAGGGGGCACGGAAGCCGACTACCAGCTGCACAA | 420 |
| Qy | 546 | CGTCCAGGTGATCTGCCACACAGAGGCGGTGGCCGAGAAGCTGGGCCAGCAGGTGAACCG | 605 |
| Db | 421 | CGTCCAGGTGATCTGCCACACAGAGGCGGTGGCCGAGAAGCTGGGCCAGCAGGTGAACCG | 480 |

[illegible]

